DOCUMENT RESUME

ED 244 964 TM 840 124

AUTHOR Nelson, Rosemery O.
TITLE Current Status and New Developments in Behavioral Assessment.
PUB DATE Dec 83
NOTE 15p.; Paper presented at the Annual Meeting of the Association for Advancement of Behavior Therapy (Washington, DC, December, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Behavioral Science Research; Behavior Modification; Clinical Psychology; *Evaluation Criteria; *Measurement Techniques; *Student Behavior; Testing Problems; Test Reliability; Test Validity
IDENTIFIERS *Behavioral Assessment

ABSTRACT

Current status and new developments in behavioral assessment for clinicians and researchers are discussed. The field of behavioral assessment has attained a recognizable identity in recent years. Behavioral assessment can be defined as the identification of meaningful response units and their controlling variables for the purposes of understanding, and altering, behavior. Practitioners have become disillusioned with the impracticality of obtaining and using behavioral assessment devices. However, there is improved accessibility of behavioral assessment devices as a result of many books and journals now available with this information. The author recommends development of a behavioral assessment kit to further improve accessibility. Researchers are disappointed with the results when psychometric criteria are used to evaluate the quality of behavioral assessment techniques. More appropriate evaluative criteria are being developed to assist researchers. These criteria include accuracy standards, individualized approaches, identification of controlling variables, and treatment validity. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED244964

Current Status and New Developments in Behavioral Assessment

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
R. O. Nelson
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Rosemery O. Nelson
Psychology Department
University of North Carolina
Greensboro, NC 27412
919-379-5013

In A. E. Kazdin (Chair), Behavioral assessment: Historical development, advances, and current status. Symposium conducted at the meeting of the Association for Advancement of Behavior Therapy, December, 1983, Washington, D. C.

NOTE: I thank Steven C. Hayes for his contribution to the content of this paper.

Related to the following papers:

Nelson, R. O. (1983). Behavioral assessment: Past, present, and future. Behavioral Assessment, 5, 195-206.

Nelson, R. O., & Hayes, S. C. (1983). Behavioral assessment: Creating a comprehensive alternative to traditional assessment approaches. Manuscript submitted for publication.

TM 840 124

Current Status and New Developments in Behavioral Assessment

My presentation this morning addresses the current status and new developments in the field of behavioral assessment.

Turning to the current status of behavioral assessment, there is some good news and some bad news. First, the good news. The field of behavioral assessment has assumed a recognizable identity during the last seven years. During this time, two journals have been founded that are devoted to advancing behavioral assessment (Behavioral Assessment published by Pergamon under the auspices of the Association for Advancement of Behavior Therapy, and Journal of Behavioral Assessment published by Plenum). Also, several books have been written that describe the techniques of behavioral assessment and their application to various clinical populations (Barlow, 1981; Ciminero, Calhoun, & Adams, 1977; Cone & Hawkins, 1977; Haynes, 1978; Haynes & Wilson, 1979; Hersen & Bellack, 1976, 1981; Keefe, Kopel, & Gordon, 1978; Mash & Terdal, 1976, 1981). As a synthesis of these efforts, behavioral assessment can be defined as the identification of meaningful response units and their controlling variables for the purposes of understanding and altering behavior (Nelson & Hayes, 1979).

Now for the bad news. There is a growing disillusionment about behavioral assessment among both practitioners and researchers. Fortunately, some new developments may stem this pessimistic tide.

Current Status and New Developments for Clinicans

Current Status. Some practitioners are currently disillusioned with the impracticality of obtaining and using behavioral assessment devices. The results of four surveys produced the following three conclusions about the assessment practices of behavior therapists. First, behavior therapists frequently use a mixture of behavioral and traditional assessment devices, including the MMPI, Rorschach, and TAT (Swan

& MacDonald, 1978; Wade & Baker, 1977; Wade, Baker, & Hartmann, 1979). Second, some behavioral assessment devices are used with greater frequency than others. Generally speaking, the only behavioral assessment techniques used with more than half of the cases seen by behavior therapists were the interview and self-monitoring (Swan & MacDonald, 1978; Wade et al., 1979). Other less frequently used behavioral assessment techniques were: behavioral observations (used with 33.9% of clients, Wade et al., 1979, and with 39.6% of clients, Swan & MacDonald, 1978); and behavioral questionnaires or self-report inventories (used with 20.3% of clients, Wade et al., 1979; and with 26.9% of clients, Swan & MacDonald, 1978). Third, behavior therapists tended to use behavioral assessment strategies rather than specific behavioral assessment techniques (Swan & MacDonald, 1979). In other words, they tended to take an individualized or idiographic approach to behavioral assessment.

Although behavior therapists generally considered assessment to be important, they reported that behavioral assessment was problematic in the following ways. First, some thought behavioral assessment was impractical in applied settings because of unfeasibility or unavailability of assessment devices (43.8% of respondents, Wade et al., mean importance rating of 2.9 on a 5-point scale, Ford & Kendall, 1979). Second, some thought behavioral assessment was inefficient or overly time-consuming (20.1% of respondents, Wade et al., 1979; mean importance rating of 3.6, Ford & Kendall, 1979). Third, some thought that lack of client cooperation interfered with behavioral assessment (mean importance rating of 3.0, Ford & Kendall, 1979).

Newer developments. Two recent developments may prove to be beneficial to clinicians who wish to use behavioral assessment strategies. The first development is as improved accessibility of behavioral assessment devices. The behavioral assessment books and journals, noted earlier, have put many devices in the public arena. There remains great room for improvement here, however, as limitations in journal space sometimes preclude publication of protocols for assessment questionnaires or interviews, or adequate descriptions of observation codes or tools for physiological

measurement. There exists a behavioral assessment underground, in which researchers share their assessment devices with each other. These devices in their entirety, however, seldom see the light of publication or find their way into the hands of practicing clinicians.

To further the accessibility of behavioral assessment devices, I would suggest that a behavioral assessment kit be developed. Such a kit might include the following items: a stopwatch and various types of observation forms that could be used while making observations in naturalistic, contrived, or role-playing situations (e.g., Marital Interaction Coding System by Hops, Wills, Patterson, & Weiss, 1971); self-monitoring forms and counting devices (e.g., Lindsley's wrist golf counter, 1968); various checklists and rating scales (e.g., Quay's Behavior Problem Checklist, 1977); various self-report questionnaires (e.g., Social Avoidance and Distress Scale, Watson & Friend, 1969), behavioral interview outlines (e.g., Holland, 1970; Kanfer & Saslow, 1969); generic role-playing instructions and samples of role-playing scenarios (e.g., Eisler, Hersen, Miller, & Blanchard, 1975, for scenes requiring assertive responses; Perri & Richards, 1979, for heterosocial scenes); and portable devices for psychophysiological measurement, such as the type used in biofeedback (Rugh & Schwitzgebel, 1975). An overview of various assessment devices suitable for clinical work is provided by Barlow, Hayes, and Nelson (1984).

Such a behavioral assessment kit, however, would have to be used <u>idiographically</u>, taking into account the unique problems and situations of each client. Another development in behavioral assessment, to be elaborated a bit later, is to recognize the <u>individuality</u> of the behavioral assessment enterprise. Clinicians need no longer conduct their idiosyncratic behavioral assessments in secret, but can acknowledge with pride that this is the proper way to perform behavioral assessment. Behavioral assessment must be implemented idiographically, so that, for each client, target behaviors can be identified, treatment strategies selected, and outcome measures determined.

To summarize then, clinicians sometimes express a disappointment with behavioral assessment, using some behavioral assessment techniques more than others and intermixing them with traditional assessment techniques. New developments which may be of assistance here are small but steady improvements in the accessibility of behavioral assessment devices and a recognition that it is desirable to use these devices in an idiographic manner.

Current Status and New Developments for Researchers

Current status. The current disillusionment with behavioral assessment is felt by researchers as well as by clinicians. For researchers, the source of this disillusionment is the unsatisfactory psychometric properties, found to be characteristic of many behavioral assessment techniques.

The most commonly advocated and utilized way to evaluate the quality of behavioral assessment techniques has been through psychometric criteria and its close relative, generalizability theory. To illustrate, Cone (1977, p. 411) wrote: "It is the thesis of the present paper that the concepts of reliability and validity used in traditional assessment are applicable to behavioral assessment as well." Similarly, Curran and Mariotto (1980, p. 33) wrote: "A refamiliarization with traditional psychometric theory combined with the utilization of generalizability theory concepts should assist us in improving our assessment procedures."

In this spirit, the psychometric properties of most behavioral assessment techniques have been tested and in many cases have been found wanting. For example, reliability across behavioral interviewers in identifying clients' problems averaged only 0.55 (Hay, Hay, Angle, & Nelson, 1979). To further illustrate, the concurrent validity among various measures of assertiveness (The College Self-Expression Scale, Galassi, DeLo, Galassi, & Bastien, 1974; the Dominance Scale of the California Psychological Inventory, Gough, 1959; peer ratings of assertiveness; and role-playing) ranged from 0.15 to 0.61 (Gorecki, Dickson, & Ritzler, 1981). Finally, the concurrent validity between a self-report measure of subjective sexual arousal and various physiological measures of sexual arousal varied from 0.39 to 0.89 (Henson, Rubin, & Henson,

1979).

While it is clear that psychometric and generalizability procedures can be applied to behavioral assessment techniques, the question is if it is appropriate to utilize these criteria for behavioral assessment. My colleague Steve Hayes and I have argued that disillusionment and disappointment is inevitable if our behavioral assessment peg continues to be forced into a traditional assessment hole (Nelson & Hayes, 1983).

We believe that psychometric criteria are an inappropriate yardstick against which to evaluate the quality of behavioral assessment techniques for the following three reasons. The first reason is that the assumptions of psychometric theory and of behavior theory differ quite radically. In psychometric theory, an observed score is a composite of the true score plus measurement error. Reliability and validity assess the consistency of measurement. Consistent data are thought to be the hallmark of a good assessment device because it is assumed that more of the true score is being measured, and less measurement error is being made. It is an assumption of behavioral assessment, however, that behavior is not necessarily enduring and consistent. To put it bluntly, there is no true behavior to measure. Inconsistency in measurement may be produced by actual changes in behavior and not by an imprecise assessment technique (Nelson, Hay & Hay, 1977). When low reliability or validity coefficients are obtained, behavioral assessors ask: "Is it the measure or the behavior?" (Cone, 1981, p. 55).

A second reason why psychometrics and generalizability theory may not be an appropriate criterion to evaluate the quality of behavioral assessment is that psychometrics are based on group data, whereas the philosophy behind behavioral assessment requires that the measure have value at the level of the individual (Cone, 1981; Nelson & Hayes, 1983). Psychometrics was built on the variability in scores produced by a group of individuals. In contrast, behaviorism is generally interested in understanding the variability in the behavior of an individual across situations and across time. General principles are sought inductively, derived from the analyses

of individuals. Variability is not seen as normal variation or as error variance, but rather variability is considered to be the product of yet-to-be-identified variables (Cone, 1981; Johnston & Pennypacker, 1980).

The third and final reason why psychometrics and generalizability theory may not be appropriate criteria to evaluate the quality of behavioral assessment strategies is that their roots lie in structuralism, whereas the roots of behaviorism lie in functionalism. As first used in psychology, structuralism examined structures of the mind such as images or sensations, whereas functionalism examined functions of the mind such as memory, imagination, and attention (Murray, 1983). In assessment, on the one hand, psychometrics generally takes a structural stance, for example, in examining how individuals' scores on two measurement devices relate to each other. On the other hand, behavioral assessment generally takes a functional stance, asking if an assessment device serves useful functions, such as selecting appropriate target behaviors or adequately measuring treatment outcome.

To summarize, then, many researchers are disappointed with the results when psychometric criteria are used to evaluate the quality of behavorial assessment techniques. These psychometric criteria, based on assumptions of true scores, based on group data, and based in structuralism, may provide an inappropriate means to evaluate the quality of behavioral assessment.

New developments. Fortunately, there are at least four new tools available to cut the psychometric bolla that are entangling the feet of behavioral assessment.

A first alternative criterion that can be used to evaluate the quality of behavioral assessment is the accuracy of measurement produced by behavioral assessment (Foster & Cone, 1980; Cone, 1981; Johnston & Pennypacker, 1980). The accuracy of an assessment device is established by comparing the measurements that it produces against some standard. In establishing the accuracy of observations made by human observers, for example, the standard might be a mechanical record of the same behavior, records made of a mechanically generated response, or a performance that has been orchestrated according to a predetermined script (Foster & Cone, 1980). The main

problem with using accuracy as a criterion to evaluate behavioral assessment is finding a suitable standard against which to evaluate accuracy. Moreover, in many situations, if such a standard exists, the standard itself may actually be the best available assessment device (Foster & Cone, 1980).

A second alternative to psychometric procedures which focus on group data are attempts to redirect behavioral assessment to the analysis of individuals. One classic idiographic assessment strategy is the functional analysis in which the controlling variables unique to the individual's responses are determined (Baer, Wolf, & Risley, 1968; Ferster, 1965; Kanfer & Saslow, 1969). The hypothesis-testing model introduced by Shapiro (1966, 1970) is another strategy to formulate and test hypotheses about individual clients. A third approach to individual assessment is the template-matching approach recently suggested by Cone (1980). In this approach, target behaviors for an individual client are identified by matching the client's current behaviors to a template of individually-determined ideal behaviors. In a fourth approach to individual assessment, idiographic correlations for a single individual among measures across time, instead of nomothetic correlations, can be computed. In our laboratory, Steve Hayes and I are currently comparing the results produced by idiographic and nomothetic correlations for sexual responding and for social skills. Already it is clear that very different results are achieved by the two approaches.

One difficulty with a completely individualized approach to assessment is that it may be very labor-intensive to develop and evaluate assessment anew for each client. A second difficulty is that an individualized approach to assessment seems to ignore the commonalities in behavior that are seen across individuals.

A third alternative way to approach the quality of behavioral assessment is to identify the measurement variables that contribute to the variability in assessment data. Assessment data vary, depending on the exact circumstances in which the assessment device is used. Therefore, generic questions about the accuracy, reliability, or validity of assessment devices are probably meaningless questions. To illustrate,

the generic question "Do self-recorders produce accurate data?" cannot be answered. The accuracy of self-recorders' data has been shown to be influenced by a host of variables, for example, training given to self-recorders , their awareness that their self-monitored data will be compared to some criterion, and reinforcement given to self-recorders for accurate data (McFall, 1977; Nelson, 1977).

The problem with this approach is which criterion to use to compare the data produced by assessment devices under different conditions. Once a criterion is established, (e.g., accuracy, treatment validity), however, then this approach allows us to determine which variables to utilize to produce effective assessment.

A fourth alternative to evaluate the quality of behavioral assessment is in terms of its contribution to treatment outcome, or in terms of "treatment validity" (Nelson & Hayes, 1979). The treatment validity question, as phrased by Mash (1979, p. 27) is: "....does the quantity and quality of assessment contribute to the effectiveness of treatment?"

Treatment validity can be determined for each aspect of behavioral assessment. To illustrate, the treatment validity effects of classifying clients according to particular target behaviors has been shown in the following study. Depressed women who displayed irrational cognitions improved more when they received cognitive therapy rather than social skills training, whereas depressed women who displayed social skills deficits improved more when they received social skills training (McKnight, Nelson, Hayes, & Jarrett, in press). The treatment validity of many other aspects of behavioral assessment remains to be tested. For example, is treatment improved when the intervention is based on a functional analysis of the individual client's problems, or is it equally effective to use a standard treatment that is generally effective with problems that match the client's topography? Does the use of multiple assessment methods lead to better treatment than the use of only one method? Does the collection of dependent measures lead to better treatment than not collecting such measures? Does the use of single subject designs improve treatment more than not using such designs?

While assessment devices with good treatment validity might also have adequate reliability, validity, or accuracy, these structural properties are neither necessary nor sufficient for treatment validity. Reliable, valid, and accurate devices are not necessarily useful. Conversely, some useful devices may not be reliable, valid, or accurate. Examples might include the clinical interview or self-report measures. Perhaps perfect accuracy is not needed for adequate utility.

Conclusion

In conclusion, despite the disappointment in behavioral assessment currently felt by clinicians and researchers alike, there is some reason for optimism. The assessment task of the clinician will be made easier as behavioral assessment devices are made more accessible and as clinicians are freed to use these devices in an idiographic manner. For the researcher, the psychometric shackles that have slowed the progress of behavioral assessment can be shed, to be replaced by more appropriate evaluative criteria, such as accuracy standards, individualized approaches, identification of controlling variables, and treatment validity. As Cicero said: "Let us not go over old ground, let us rather prepare for what is to come."

References

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 91-97.

Barlow, D. H. (Ed.). (1981). Behavioral assessment of adult disorders. New York: Guilford.

Barlow, D. H. (1983), Hayes, S. C., & Nelson, R. O. (1984). The scientist-practitioner: Research and accountability in clinical and educational settings. New York: Pergamon.

Ciminero, A. R., Calhoun, K. S., & Adams, H. E. (Eds.). (1977). Handbook of behavioral assessment. New York: Wiley.

Cone, J. D. (1977). The relevance of reliability and validity for behavioral assessment. Behavior Therapy, 8, 411-426.

Cone, J. D. (1980, November). Template matching procedures for idiographic behavioral assessment. Paper presented at the meeting of the Association for Advancement of Behavior Therapy, New York.

Cone, J. D. (1981). Psychometric considerations. In M. Hersen & A. S. Bellack (Eds.), Behavioral assessment. (pp. 38-68). New York: Pergamon.

Cone, J. D., & Hawkins, R. P. (Eds.). (1977). Behavioral assessment: New directions in clinical psychology. New York: Brunner/Mazel.

Curran, J. P., & Mariotto, M. J. (1980). A conceptual structure for the assessment of social skills. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), Progress in behavior modification (Vol. 10, pp. 1-37). New York: Academic.

Eisler, R. M., Hersen, M., Miller, P. M., & Blanchard, E. B. (1975). Situational determinants of assertive behaviors. Journal of Consulting and Clinical Psychology, 43, 330-340.

Ferster, C. B. (1965). Classification of behavioral pathology. In L. Krasner & L. P. Ullmann (Eds.) Research in behavior modification. (pp. 6-26). New York: Holt, Rinehart, & Winston.

Ford, J. D., & Kendall, P. C. (1979). Behavior therapists' professional behaviors: Converging evidence of a gap between theory and practice. The Behavior Therapist, 2 (5), 37-38.

Foster, S. L., & Cone, J. D. (1980). Current issues in direct observation. Behavioral Assessment, 2, 313-338.

Galassi, J. O., DeLo, J. S., Galassi, M. D., & Bastien, S. The College Self-Expression Scale: A measure of assertiveness. Behavior Therapy, 5, 165-171.

Gorecki, P. R., Dickson, A. L., & Ritzler, B. A. Convergent and concurrent validation of four measures of assertion. Journal of Behavioral Assessment, 3, 85-91.

Gough, H. G. (1957). California Psychological Inventory Manual. Palo Alto: Consulting Psychologists Press.

Hay, W. M., Hay, L. R., Angle, H. V., & Nelson, R. O. (1979). The reliability of problem identification in the behavioral interview. Behavioral Assessment, 1, 107-118.

Haynes, S. N. (1978). Principles of behavioral assessment. New York: Gardner.

Haynes, S. N., & Wilson, C. C. (1979). Behavioral assessment: Recent advances in methods, concepts, and applications. San Francisco: Jossey-Bass.

Henson, D. E., Rubin, H. B., & Henson, C. (1979). Analysis of the consistency of objective measures of sexual arousal in women. Journal of Applied Behavior Analysis, 12, 701-711.

Hersen, M., & Bellack, A. S. (Eds.) (1976). Behavioral assessment: A practical handbook. New York: Pergamon.

Hersen, M., & Bellack, A. S. (Ed.) (1981). Behavioral assessment: A practical handbook (2nd ed.). New York: Pergamon.

Holland, C. J. (1970). An interview guide for behavioural counseling with parents. Behavior Therapy, 1, 70-79.

Hops, H., Wills, T. A., Patterson, G. R., & Weiss, R. L. (1971). Marital interaction coding system. University of Oregon and Oregon Research Institute. Order from ASIS? NAPS, c/o Microfiche Publications, 305 E. 46th Street; New York, NY 10017.

Johnston, J. M., & Pennypacker, H. S. (1980). Strategies and tactics of human behavioral research. Hillsdale, N. J.: Erlbaum.

Kanfer, F. H., & Saslow, G. (1969). Behavioral diagnosis. In C. M. Franks (Ed.), Behavior therapy: Appraisal and status, (pp. 417-444). New York: McGraw-Hill.

Keefe, F. J., Kopel, S. A., & Gordon, S. B. (1978). A practical guide to behavioral assessment. New York: Springer.

Lindsley, O. R. (1968). A reliable wrist counter for recording behavior rates. Journal of Applied Behavior Analysis, 1, 77-78.

Mash, E. J. (1979). What is behavioral assessment? Behavioral Assessment, 1, 23-29.

Mash, E. J., & Terdal, L. G. (Eds.). (1976). Behavior therapy assessment. New York: Springer.

Mash, E., & Terdal, L. (Eds.), (1981). Behavioral assessment of childhood disorders. New York: Guilford.

McFall, R. M. (1977). Parameters of self-monitoring. In R. B. Stuart (Ed.), Behavioral self-management: Strategies, techniques, and outcomes (pp. 196-214). New York: Brunner/Maxel.

McKnight, D. L., Nelson, R. O., Hayes, S. C., & Jarrett, R. B. (in press). Importance of treating individually-assessed response classes in the amelioration of depression. Behavior Therapy.

Murray, D. J. (1983). A history of Western psychology. Englewood Cliffs, N. J.: Prentice-Hall.

Nelson, R. O. (1977). Assessment and therapeutic functions of self-monitoring. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), Progress in behavior modification Vol. 5, pp. 263-308). New York: Academic.

Nelson, R. O., Hay, L. R., & Hay, W. M. (1977). Comments on Cone's "The relevance of reliability and validity for behavioral assessment." Behavior Therapy, 8, 427-430.

Nelson, R. O., & Hayes, S. C. (1979). Some current dimensions of behavioral assessment. Behavioral Assessment, 1, 1-16.

Nelson, R. O., & Hayes, S. C. (1983). Behavioral assessment: Creating a comprehensive alternative to traditional assessment approaches. Manuscript submitted for publication.

Perri, M. G., & Richards, C. S. (1979). Assessment of heterosocial skills in male college students. Behavior Modification, 3, 337-354.

Quay, H. C. (1977). Measuring dimensions of deviant behavior: The Behavior Problem Checklist. Journal of Abnormal Child Psychology, 5, 277-289.

Rugh, J. D., & Schwitzgebel, R. L. (1975). Biofeedback apparatus: List of suppliers. Behavior Therapy, 6, 238-240.

Shapiro, M. B. (1966). The single case in clinical-psychological research. The Journal of General Psychology, 74, 3-23.

Shapiro, M. B. (1970). Intensive assessment of the single case: An inductive-deductive approach. In P. Mittler (Ed.), The psychological assessment of mental and physical handicaps, (pp. 645-666). London: Methuen.

Swan, G. E., & MacDonald, M. L. (1978). Behavior Therapy, 9, 799-807.

Wade, T. C., & Baker, T. B., (1977). Opinions and use of psychological tests: A survey of clinical psychologists. American Psychologist, 32, 874-882.

Wade, T. C., Baker, T. B., & Hartmann, D. P. (1979). Behavior therapists' self-reported views and practices. The Behavior Therapist, 2 (1), 3-6.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. Journal of Consulting and Clinical Psychology, 33, 448-457.